



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

LINEAR ALGEBRA
AND ITS
APPLICATIONS

Linear Algebra and its Applications 388 (2004) 173–191

www.elsevier.com/locate/laa

Linearly independent set families

Friedhelm Eicker ^{a,*}, Guenter Ewald ^b^a*FB Statistik, Dortmund-University, D44221 Dortmund, Germany*^b*FB Mathematik, Ruhr-University, D44801 Bochum, Germany*

Received 18 February 2002; accepted 19 February 2004

Submitted by G.P.H. Styan

Abstract

An apparently new definition of linearly independent set families in a linear space R^k is given (for short ‘independent set families’) and a relation of such families to families of separating hyperplanes is established. Independent set families are families of k subsets of the R^k defined by the property that any k points P_i , one from each subset, are linearly independent. The concept is not related to that of independent subsets of a finite basic set used in matroid theory. A typical example of an independent set family arising from a statistical application consists of congruent narrow cones around the unit vectors u_i of the R^k which are open and convex. The statistical application referred to is robust multilinear regression.

© 2004 Elsevier Inc. All rights reserved.

AMS classification: 15A03; 62J05*Keywords:* Linearly independent set families; Independent set families; Separating hyperplanes; Robust multilinear regression

1. Introduction

At the core of the paper is an apparently new definition (Definition 1) of linearly independent set families (for short ‘independent set families’) in linear spaces and their relation to families of separating hyperplanes (Theorem 2.3). Independent set families are families of k subsets of the R^k defined by the property that any k points P_i , one from each subset, are linearly independent. A typical example of an independent set family which arises from a statistical application consists of narrow

* Corresponding author. Tel.: +49-321-759832.

E-mail address: friedhelm.eicker@udo.de (F. Eicker).

congruent cones around the unit vectors u_i of the R^k which are open and convex (Theorem 3.1). The cones grow increasingly narrower as the dimension k increases. A selection of k representative points from them consists of perturbed unit-vectors; they can be written as a $k \times k$ -matrix which is non-singular. Further examples are given in the text. There exist covering theorems for the R^k by sets of at most $k + 1$ independent set families but these are not treated in this paper. The concept of an independent set family used here has nothing to do with the concept of independent subsets of a finite basic set used in the matroid definition [4]; compare also the last paragraph of Section 2.

The statistical application which gave rise to this paper is robust multilinear regression outlined in [1,2].

2. Independent set families and separating hyperplanes

2.1. Definition and discussion of independent set families

The definition of independent set families extends in a straightforward way the definition of linearly independent points to independent sets in a linear space. The formal definition is

Definition 1. A family of non-empty subsets $C_i \subset R^k$

$$\mathcal{F} := \{C_i, i = 1, \dots, k\}$$

is called a *linearly independent set family* (for short: an ‘independent set family’) if and only if any selection of k points $P_i \in C_i$ is linearly independent.

Obviously a family of k linearly independent (non-collinear) points in the usual sense is the simplest example (so the definition is not empty).

Remark 2.1. If $\{C_1, \dots, C_k\}$ is an independent set family then

- (i) The C_i are disjoint and none of them includes the origin.
- (ii) $\{C_i^*, i = 1, \dots, k\}$ is an independent set family for any choice of non-empty subsets $C_i^* \subset C_i$.

A family $\{C_1, \dots, C_k\}$ is linearly independent if and only if one of the following conditions is true:

- (iii) There does not exist a proper linear subspace H of R^k which contains points from two or more of the C_i .
- (iv) The family $\{\lambda_i C_i, i = 1, \dots, k\}$ is linearly independent for any coefficients $\lambda_i \neq 0$.
- (v) The family $\{MC_i, i = 1, \dots, k\}$ is an independent set family for any non-degenerate linear transformation M of the R^k .

- (vi) $\{\lambda P : P \in C_i, \text{ any } \lambda \neq 0\}, i = 1, \dots, k$, is an independent set family. Its sets consist of ‘pinched’ lines that pass through a point of C_i and through 0 without containing 0. Hence these sets are pinched cones; they will be considered in more detail below.
- (vii) The family $\{C_i \cup (-C_i), i = 1, \dots, k\}$ is an independent set family.

The following lemma gives an illustration of the interplay between linear independence and convexity for simple point sets that later-on is generalized to general point sets (Theorem 2.3(ii)).

Lemma 2.1. *Let $\{P_1, \dots, P_k\}$ and $\{P_1, \dots, P_{k-1}, P'_k\}$, $P_k \neq P'_k$, be two linearly independent sets of points in the R^k . Then each point P_k and P'_k lies in an open half-space associated with the hyperplane*

$$H := \text{lin}(P_1, \dots, P_{k-1}), \quad (2.1)$$

the linear space spanned by the points P_1, \dots, P_{k-1} . They lie in different half-spaces if and only if there exists a real number $\lambda \in (0, 1)$ such that the points $P_1, \dots, P_{k-1}, \lambda P_k + (1 - \lambda)P'_k$ are linearly dependent. They lie in the same half-space if and only if the line segment

$$S := \{\lambda P_k + (1 - \lambda)P'_k : \lambda \in [0, 1]\}$$

lies entirely in that half-space, and then the family of the subsets

$$\{\{P_1\}, \dots, \{P_{k-1}\}, S\}$$

is an independent set family.

Proof. With the representation $H(x) = \{n'x = 0\}$ (n a suitable coefficient vector in R^k) the scalar products $n'P_k$ and $n'P'_k$ are not null under the assumption which implies that P_k and P'_k lie in the same or in different open half-space(s) of the hyperplane H . We use

$$H(\lambda P_k + (1 - \lambda)P'_k) = \lambda n'P_k + (1 - \lambda)n'P'_k.$$

The points P_k and P'_k lie in the same half-space if and only if $n'P_k$ and $n'P'_k$ have the same sign or, equivalently, all points $H(\lambda P_k + (1 - \lambda)P'_k), \lambda \in [0, 1]$, have that sign.

The signs of $n'P_k$ and $n'P'_k$ are different if and only if the linear function $[\text{in } \lambda] \lambda n'P_k + (1 - \lambda)n'P'_k$ has exactly one zero inside $[0, 1]$. \square

Example

- (i) In R^2 consider the open first and fourth quadrants

$$C_1 := \{(x_1, x_2) : x_1 > 0, x_2 > 0\}, \quad C_2 := \{(x_1, x_2) : x_1 < 0, x_2 > 0\}.$$

Then $\mathcal{F} := \{C_1, C_2\}$ is an independent set family, and so is

$$\mathcal{F}' := \{C_1 \cup (-C_1), C_2 \cup (-C_2)\}.$$

The set C_1 may be replaced by one of the half-open sets

$$\{(x_1, x_2) : x_1 \geq 0, x_2 > 0\} \text{ or } \{(x_1, x_2) : x_1 > 0, x_2 \geq 0\}$$

or by the set $\{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\} \setminus \{0\}$ which is closed except at 0. Similar replacements can be made with C_2 but not all replacements can be made with both sets at the same time.

(ii) Again in R^2 consider the parallel to the vertical axis $\{(1, x_2) : x_2 \in R^1\}$ and in it the subsets

$$C_1 := \{(1, x_2) : x_2 \text{ rational}\}, \quad C_2 := \{(1, x_2) : x_2 \text{ irrational}\}.$$

Then $\mathcal{F} := \{C_1, C_2\}$ is an independent set family.

(iii) Let L^i be i -dimensional linear subspaces of R^k , $i = 1, \dots, k$, $L^0 := \{0\}$, and let $L^0 \subset L^1 \subset \dots \subset L^k$. Then the ‘flag’

$$\{L^i \setminus L^{i-1} : i = 1, \dots, k\}$$

is an independent set family.

We now recall some general notions from convex and general geometry that are needed subsequently.

If C is a subset of R^k , then $\text{lin } C$ denotes the *smallest linear subspace* of R^k containing C .

A (linear) *hyperplane* of R^k is a (linear) $(k-1)$ -dimensional subspace of R^k and is denoted by $H = \{x : \langle x, a \rangle = 0\}$ with some non-zero coefficient vector $a \in R_k$; here $\langle \cdot, \cdot \rangle$ denotes the ordinary inner product of R^k . We often write $\langle x, a \rangle = x'a =: H(x)$ for the associated linear operator. Without the adjective ‘linear’ we mean the linear hyperplane as just defined. The definition of half-spaces bounded by a hyperplane is postponed until after Definition 4.

We shall use the operators (for any non-empty set $M \subset R^k$, also called *generating set*)

$$\begin{aligned} \text{conv } M &:= \{\alpha_1 a_1 + \dots + \alpha_k a_k : a_i \in M, 0 \leq \alpha_i, i = 1, \dots, k, \\ &\quad \alpha_1 + \dots + \alpha_k = 1 \text{ for all } k \geq 1\}, \\ \text{pos } M &:= \{\alpha_1 a_1 + \dots + \alpha_k a_k : a_i \in M, 0 \leq \alpha_i, i = 1, \dots, k, k \geq 1\}, \\ \text{cl } M &:= M \cup \{\text{cluster points of } M\}. \end{aligned}$$

In this paper we use the definition of cones common in convex geometry:

Definition 2 [6]. A set C is called a (linear) *cone* and 0 is called an *apex of the cone* if $x \in C$ and $\lambda \geq 0$ imply $\lambda x \in C$.

The set $C \cup (-C)$ is called the *extended (linear) cone* associated with C . It is a union of lines whereas a cone is a union of rays.

A *pinched cone* is obtained from a cone by taking away the apex (the origin in the above case).

An ‘open cone’ is defined as the open kernel (if $\neq \emptyset$) of a convex cone. A ‘polyhedral cone’ is defined as the intersection of a finite number of half-spaces.

From linear hyperplanes, half-spaces, cones, extended cones, and polyhedral cones we proceed to the corresponding *affine* notions by adding the translation vector a . Without the adjective ‘linear’ we mean the linear notions, whereas we add ‘affine’ in the affine case. If C is a linear cone, then a is an *affine apex* of the *affine cone* $C + a$. An affine cone may have many apices. For example, each point $b \in H + a$, an affine hyperplane, is an apex of the affine half-space $H^+ + a$, and R^k is an affine cone consisting of apices only.

If the intersection of a finite number of affine half-spaces is bounded and has interior points we call it a *k-polytope*, in particular, a *k-simplex* if the number of half-spaces is $k + 1$. An affine subspace A of R^k may be considered a linear space (by choosing 0 on it), so relative to A we may define an *m-polytope* if m is the dimension of A . If an affine polyhedral cone $C + a$ is given which has a as its only apex, it can readily be shown that there exists an affine hyperplane K such that $K \cap (C + a)$ is a polytope.

Definition 3 (*Mapping onto a $(k - 1)$ -dimensional surface or onto an affine hyperplane*). Let a $(k - 1)$ -dimensional surface W surrounding (but not containing) the origin 0 be given in such a way that every ray hits W exactly once. Any point $P \in R^k \setminus \{0\}$ is then mapped uniquely into a point $P' \in W$ (thus defining a central projection). This mapping extends to subsets: any $S \subset R^k \setminus \{0\}$ is mapped into a set $S' \subset W$. If a pinched cone C is mapped in this way onto its image in $D \subset W$ we say that C is *generated* by D and we write $C := \text{cone } D = \text{pos } D \setminus \{0\}$. Similarly, given an affine hyperplane $H + a$ a unique mapping of all points $P \in R^k \setminus H$ onto $H + a$ can be defined using lines passing through 0 and P instead of rays.

Examples of such surfaces W are centered spheres, k -dimensional simplices or cubes.

Remark 2.2. For the set of all independent set families there exist a number of equivalences as follows. Let $\{D_i\}$, $i = 1, \dots, k$, be a family of subsets and let the centrally projected images on W be $\{D'_i \subset W\}$ and the associated pinched cones *cone* D'_i . If one of these three families is an independent set family then so are the other two.

2.2. Separating hyperplanes and independent set families

The main result of this section is a characterization of independent convex sets by separating hyperplanes (Theorem 2.3). We need the standard definition (compare e.g. [5]).

Definition 4

- (i) Let f be a real-valued linear functional on R^k and let $A \subset R^k$ be a non-empty subset. Then

$f(A) \geq \alpha$ means that $f(x) \geq \alpha$ for all $x \in A$ and some real α .

Similar definitions hold for the reversed or strict inequalities.

An open half-space generated by the affine hyperplane $h_f := \{x : f(x) = \alpha\}$ is defined by

$$\{x \in R^k : f(x) > \alpha\} =: (h_f^+)^o \text{ respectively by } \{x \in R^k : f(x) < \alpha\} =: (h_f^-)^o$$

where the superscript o indicates openness.

- (ii) The hyperplane h_f separates in the wide sense the non-empty subsets A and B of R^k if either $f(A) \geq \alpha$, $f(B) \leq \alpha$ or the inverse of both inequalities holds.
 (iii) The hyperplane h_f separates properly the two sets if and only if strict inequalities hold. For short in this case we also just say that h_f separates the two sets.

If the linear functional $f(x)$ is defined using the inner product $f(x) := \langle x, a \rangle$ (with a given non-zero $a \in R^k$) then we set $H^+ := \{x : \langle x, a \rangle \geq 0\}$ and $H^- := \{x : \langle x, a \rangle \leq 0\}$, called the *half-spaces* bounded by the linear hyperplane $H := \{x : \langle x, a \rangle = 0\}$. If we want to interchange these half-spaces we have to use $-f(x)$ instead of $f(x)$ [respectively $-a$ instead of a]. The half-spaces are closed sets and both contain the origin as well as H . The corresponding open half-spaces are denoted, e.g., by $(H^+)^o$. However, if H is given initially as a $(k-1)$ -dimensional linear space it is up to us how to define H^+ , H^- . If we want to define them using an inner product as above the definition is determined by the choice of $-a$ or a where a in the inner product is any non-zero vector orthogonal to H . This distinction is important under certain linear transformations as used, e.g., in Lemma 2.5(iii).

We can use the separating hyperplanes of Definition 4 to characterize large classes of independent set families. The special case of k linearly independent points in R^k is rather trivial but nevertheless illustrative and useful to introduce some notation helpful also in the general case:

Notation. Let \mathcal{T} be the family of all subsets $S \subset \{1, \dots, k\}$ with the exception that of S and S^c only that set belongs to \mathcal{T} which contains the larger natural number; hence e.g. $S = \emptyset$ is not contained in \mathcal{T} . There are 2^{k-1} sets in \mathcal{T} .

Definition 5. Let $\mathcal{F} := \{D_i : i = 1, \dots, k\}$ be a family of non-empty sets in R^k . Any family $\mathcal{H} := \{H_S : S \in \mathcal{T}\}$ of hyperplanes H_S is called a *family of separating hyperplanes associated with the set family* $\{D_i\}$ (for short ‘an associated hyperplane-family’) if for every subfamily

$$\mathcal{F}_S := \{D_i : i \in S\} \subset \mathcal{F} \quad (2.2)$$

(i.e. for any $S \in \mathcal{T}$) there exists a linear separating hyperplane H_S which separates properly the D_i —sets of \mathcal{F}_S from the sets of the complementary family \mathcal{F}_{S^c} . For short we also say that ‘ H_S separates \mathcal{F}_S and \mathcal{F}_{S^c} ’. The family \mathcal{F}_{S^c} is empty if the complement $S^c = \emptyset$; in that case the separating hyperplane $H_{\{1, \dots, k\}}$ is a bounding hyperplane for \mathcal{F} : all D_i lie in one and the same half-space.

Perhaps the simplest example for the use of an associated family of separating hyperplanes is described in the following lemma for the family $\{D_i := \{u_i\}, i = 1, \dots, k\}$ where u_i is the i th unit-vector:

Lemma 2.2. *For the vectors $u_i, i = 1, \dots, k$, an associated family of separating hyperplanes is given by*

$$H_S := \left\{ x \in \mathbb{R}^k : \sum_{i \in S} x_i - \sum_{i \in S^c} x_i = 0 \right\}, \quad S \in \mathcal{T}. \quad (2.3)$$

Putting $n_S = \sum_{i=1}^k (1_{i \in S} - 1_{i \notin S})u_i$, a sign vector, H_S is determined by the inner product $n'_S x$.

The set of all vectors $\pm n_S$ is the set of all diagonal vectors (rays) of the 2^k quadrants (‘cells’) generated by the k coordinate hyperplanes of the \mathbb{R}^k .

The k hyperplanes $H_{\{i\}}$ for $S = \{i\}, i = 1, \dots, k$, have linearly independent coefficient vectors. $H_{\{i\}}$ separates (‘isolates’) $\{u_i\}$ from the other sets $\{u_j\}$.

The example of this lemma is generalized in Lemma 2.5.

Proof. The assertion follows from $n'_S u_i = 1_{i \in S} - 1_{i \notin S}$. \square

The example of the lemma clearly shows the interplay of geometry, combinatorics and linear algebra as the three essential and indispensable pillars of our approach. This carries over also to the case of general point sets.

In this case of general independent point set families (instead of singletons) we should expect to need infinitely many hyperplanes. However, our basic Theorem 2.3 shows that, at least for open convex cones D_i , only 2^{k-1} hyperplanes satisfying certain assumptions are needed in order to establish independence.

Theorem 2.3

- (i) A given family $\mathcal{F} = \{D_i : i = 1, \dots, k\}$ of non-empty subsets of the \mathbb{R}^k is an independent set family if there exists at least one family \mathcal{H} of separating hyperplanes associated with \mathcal{F} in the sense of Definition 5.
- (ii) If the D_i are convex and linearly independent sets then necessarily there exists an associated system \mathcal{H} of separating hyperplanes.
- (iii) Let \mathcal{F} be an independent set family and let there exist at least one associated system \mathcal{H} of separating hyperplanes with the property that its hyperplanes H_S all separate properly. Then the H_S are all different.

The intersection of all those open half-spaces of the H_S that contain a given D_i is an open polyhedral cone, any $i = 1, \dots, k$.

Definition 6. In the case (iii) of the preceding theorem the open polyhedral cones are called the *maximal cones* with respect to the chosen system \mathcal{H} .

A notational warning: In Definition 4 hyperplanes defined by a functional f carried this f as a subscript while in the above theorem the index is a set S .

Note that any cone being considered in (iii) of the above theorem may have up to 2^{k-1} sides.

Before we prove Theorem 2.3 we state a corollary of that theorem which is a weaker but more handy result:

Corollary 2.4. Let $\mathcal{F} := \{C_i, i = 1, \dots, k\}$ be a family of convex and open subsets of the R^k . Then \mathcal{F} is an independent set family if and only if an associated system of separating hyperplanes exists.

Proof (Proof of Theorem 2.3).

(1) Proof of (i) (the ‘If’-direction). The assumption is not empty because of Lemma 2.2. Assume then that there exists a system \mathcal{H} satisfying (i) for the given D_i -sets. We have to show that $\sum_{i=1}^k m_i P_i \neq 0$ for all non-zero vectors $m \in R^k$ and any selection of points $P_i \in D_i$, $i = 1, \dots, k$ (without loss of generality we assume the existence of an $m_i > 0$). We will show that there exists a suitable H_S such that $\sum_{i=1}^k m_i P_i \notin H_S$ and then infer $\sum_{i=1}^k m_i P_i \neq 0$. Let $H_S = \{x \in R^k : n'_S x = 0\}$ with a suitable a coefficient vector n_S . Consider $S := \{i : m_i > 0\}$ ($\neq \emptyset$ without loss of generality). Then in

$$n'_S \left(\sum_{i=1}^k m_i P_i \right) = \sum_{i=1}^k m_i n'_S P_i = \sum_S + \sum_{S^c} \quad (2.4)$$

the first sum on the right is positive and the latter one is non-negative due to proper separation of all D_i (which implies, e.g. $n'_S P > 0$ for all $P \in D_i \subset (H_i)^+$, any $i \in S$).

(2) Proof of (ii) (the ‘only-if’ direction).

Assume now the D_i are linearly independent and convex. Then none of them contains the origin and they are disjoint by Remark 2.1(i). We want to apply the Hahn-Banach separation theorem (compare Theorem 4.3): For every non-empty proper subset $S \subset \{1, \dots, k\}$ we show first that the convex hulls $\text{conv}(\cup_{i \in S} D_i) =: K_S$ and $\text{conv}(\cup_{i \notin S} D_i) =: K_{S^c}$ are disjoint and neither one contains the origin. Assume to the contrary that there exists an $x \in K_S \cap K_{S^c}$. By Lemma 4.1 any point $x \in K_S$ can be written as a (non-trivial) convex combination of elements of the D_i , $i \in S$, with at

most one element from each D_i . Similarly since $x \in K_{S^c}$, x can also be written as a convex combination of elements in D_i , $i \notin S$ with at most one element from each such D_i . Equalling both representations results in a vanishing non-trivial linear combination of elements from different D_i . This however contradicts the independence property. By the same argument the origin cannot be an element of K_S or of K_{S^c} .

Consequently by Theorem 4.3 and due to the disjointness, non-emptiness and convexity of K_S and K_{S^c} there exists a hyperplane H_S , in general affine, separating the two in the sense $K_S \subset H_S^+$ and $K_{S^c} \subset H_S^-$ (or with $+$ and $-$ exchanged).

For the case $S = \{1, \dots, k\}$ the existence of $H_{\{1, \dots, k\}}$ satisfying (i) may be proven using the following fact: A family D_1, \dots, D_k is an independent set family if and only if $-D_1, D_2, \dots, D_k$ is one (Remark 2.1(iii)). Consequently there exists a separating hyperplane H isolating $-D_1$ from D_2, \dots, D_k by the preceding part of the proof. Now reflection of $-D_1$ at the origin brings the reflected $-D_1$ (which is D_1) to the same side of H where D_2, \dots, D_k are. Consequently this hyperplane H may be taken as $H_{\{1, \dots, k\}}$.

The separating hyperplanes can be chosen so as to contain the origin since trivially the D_i can be embedded into pinched convex cones K_i generated by the D_i and with apex 0 (but not containing 0), and these cones also form a family of independent sets. Separating hyperplanes for them exist by the preceding steps of the proof. These hyperplanes by construction also separate the D_i and all must contain the origin since every (k -dimensional) neighborhood of the origin contains points from each cone K_i , hence separation is only possible if each separating hyperplane contains 0 (just let the radius of ball-shaped neighborhoods tend to zero). Since this is true for every i the proof is complete.

(3) Proof of statement (iii).

In order to prove that k -dimensionality (esp. openness) of all sets D_i implies the distinctness of the hyperplanes H_S consider two different sets S and S' . Let e.g. $i \in S^c \cap S'$. Then due to k -dimensionality there exists an element $P \in D_i$ such that $P \in (H_{S'}^+)^0$ (say) while also $P \in H_S^-$. The assumption $H_S = H_{S'}$ therefore yields a contradiction.

The open kernels of the cones K_i are an independent set family since any system of separating hyperplanes satisfying (i) with respect to the family $\{D_i\}$ satisfies (i) also with respect to the $\{K_i\}$ -family. \square

Remark 2.3

- (i) All the 2^{k-1} index-sets $S \in \mathcal{T}$ occurring in Theorem 2.3 are required in the assertion of that theorem. This is true since the parameter vector m occurring in step 1) of the proof of Theorem 2.3 may have only non-zero components which are all of the same sign.
- (ii) For some or most applications (such as the statistical application to robust multilinear regression) it suffices to consider open cones only in Theorem 2.3. This relieves us from the consideration of independent sets having points in

common with [wide sense] separating hyperplanes which may have to be taken into account in more general setups.

- (iii) Most extended cones are not convex but in case of ordinary cones the transition from cones to extended cones is a trivial matter and so are the related independence statements.
- (iv) Presumably the k -dimensionality of the C_i is even necessary in order that any related system of properly separating hyperplanes H_S consists of different hyperplanes only.

A proof of the remark which is not spelled out here in detail may use the following idea: Take two different sets S and S' and assume $H_S = H_{S'}$. Then by assumption there exists without loss of generality an $i \in S \cap (S')^c$ [or vice versa] such that $D_i \subset H_S^+$, which leads to a contradiction.

The following lemma and the subsequent theorem offer a construction method presumably for all independent set families of open polyhedral cones. An outline of the method is as follows: Start with an arbitrary set of non-collinear points P_i (e.g. the basic unit vectors), select a set of separating hyperplanes for them that don't contain any of the P_i and then take the open cones K_i containing P_i and being generated by the hyperplanes according to Theorem 2.3(i). Then any non-singular linear transformation can be applied yielding again an independent set family.

The simple example of Lemma 2.2 is generalized by

Lemma 2.5

- (i) For any set $\{v_1, \dots, v_k\}$ of orthogonal (and thus linearly independent) vectors v_i the singletons $D_i := \{v_i\}$ are separated by hyperplanes H_S defined via the coefficient vectors

$$h_S := \sum_{i=1}^k (1_{i \in S} - 1_{i \in S^c}) v_i.$$

- (ii) Any system of separating hyperplanes $\{H_S = \{x : x' m_S = 0\}, S \in \mathcal{T}\}$ associated with the family $\{u_i, i = 1, \dots, k\}$ is characterized by vectors $m_S = (m_{S,1}, \dots, m_{S,k})'$ satisfying

$$m_{S,i} \begin{cases} \geq 0 & \text{for } i \in S, \\ > 0 & \text{for at least one } i =: i_S \in S, \\ \leq 0 & \text{for } i \in S^c. \end{cases} \quad (2.5)$$

(The class of these vectors contains the family

$$\left\{ n_S = \sum_{i=1}^k (1_{i \in S} - 1_{i \notin S}) u_i, \quad S \in \mathcal{T} \right\},$$

which occur in Lemma 2.2, as a special case.)

- (iii) Case (ii) can be linearly transformed to any family $\mathcal{F} := \{\{P_i\}, i = 1, \dots, k\}$ of independent points P_i . Put $P := (P_1, \dots, P_k)$ (a non-singular $k \times k$ -matrix).

Then the coefficient vectors h_S of the hyperplanes H_S ($S \in \mathcal{T}$) associated with \mathcal{F} are determined by the transformation

$$h_S = (P')^{-1} m_S, \quad (2.6)$$

m_S subject to (2.5).

- (iv) Let K_i be the maximal cone, that contains P_i , $i = 1, \dots, k$ according to Definition 6. The independent points P_i (in the special example the vectors u_i) are interior points of K_i , respectively, if and only if strict inequalities hold throughout the conditions occurring in (2.5) (the middle inequality in these conditions is then automatically satisfied).

Proof

- (i) follows from the relations $h'_S v_i > 0$ resp. < 0 according to $i \in S$ or not.

Hence all v_i with $i \in S$ lie in H_S^+ , the remainder in H_S^- .

- (ii) is obvious.

- (iii) follows from $h'_S P_i = m'_S P^{-1} P_i = m'_S u_i$ and (2.5) of (ii). \square

Using Theorem 2.3(i), last sentence, we get:

Theorem 2.6. Let a family of separating hyperplanes $\{H_S, S \in \mathcal{T}\}$, $H_S = \{x : x'h_S = 0\}$, be associated with a given set of independent points $\{P_i, i = 1, \dots, k\}$. Let the h_S satisfy $h_S = (P')^{-1} m_S$ where the vectors m_S are subject to

$$m_{S,i} \begin{cases} > 0 & \text{for } i \in S, \\ < 0 & \text{for } i \in S^c. \end{cases} \quad (2.7)$$

Then the open kernels of the intersections of the half-spaces belonging to the hyperplanes H_S that contain the P_i provide a family of independent open and polyhedral cones, and the H_S separate them. Moreover the classes of all independent families of open and polyhedral cones and of all systems of hyperplanes separating them are obtained this way.

The great generality in selecting the m_S -vectors is quite useful in the statistical application that gave rise to this paper.

The situation may be seen also from a somewhat different angle. The question arises to characterize in a different way a system of associated separating hyperplanes, which has 2^{k-1} members, within the class of all families consisting of 2^{k-1} hyperplanes. Equivalently, the question may be asked to tell whether or not for a given family of 2^{k-1} hyperplanes k linearly independent vectors $P_i \in R^k$ can be found such that the family is a family of separating hyperplanes associated with them. Certainly, the requirement of strict separation should be imposed in order to make the problem more precise. Similarly, the problem is reducible if, e.g., linearly dependent points are also admitted. Anyway, here seems to lie an interesting unsolved problem a solution of which would be of interest also in other contexts.

A family certainly is a system of separating hyperplanes if there exist k cones among all the cones (cells) generated by the hyperplanes such that each one contains one of k linearly independent points P_i as interior points and with respect to which all hyperplanes of the family can be represented in the form described in Theorem 2.6. The case that some of the linearly independent points P_i lie *on* some of the separating hyperplanes is somewhat more complicated but seems tractable.

An important class of independent set families not explicitly covered by this paper consists of families of non-convex independent sets, such as extended independent cones. Some such families however can be obtained in a trivial way from families of independent convex sets.

Although certain concepts occurring in the present work as well as in (oriented) matroid theory (like families of oriented hyperplanes, cells and sign vectors; compare e.g. [8]) results of the latter theory apparently cannot be brought to bear here because, e.g., of the use of the Hahn-Banach theorem in the proof of Theorem 2.3.

3. Further examples of independent set families

A fairly simple example are rotational cones as independent subsets. They can be easily visualized in R^3 .

3.1. Congruent rotational cones

Theorem 3.1

- (i) In R^k let $u_i = (0, \dots, 0, 1, 0, \dots, 0)'$ be the i th unit-vector and $K_i, i = 1, \dots, k$, any vectors. Then the open sets

$$D_i := \{q_i = u_i + K_i : \|K_i\| < 1/\sqrt{k}\}, \quad i = 1, \dots, k, \quad (3.1)$$

provide an independent set family. The value $1/\sqrt{k}$ of the radii of these congruent balls cannot be increased if the independence property of the family is to be maintained.

Similarly, the extended open rotational cones D'_i in R^k generated by the sets D_i :

$$D'_i = \{\lambda q_i : q_i \in D_i, \lambda \neq 0\}, \quad i = 1, \dots, k; \quad (3.2)$$

are an independent set family; the D'_i have apices at the origin and central axes u_i .

- (ii) The separating tangent hyperplane $x'u = 0, u := (1, \dots, 1)'$ touches all sets of the two families defined in (i).

The independence of the family D_i is equivalent to the non-singularity of all the perturbed $k \times k$ -unit matrices $(u_1 + K_1, \dots, u_k + K_k)$ where the vectors K_i represent small perturbations with $\|K_i\| < 1/\sqrt{k}$.

The proof of the above theorem follows from

Lemma 3.2. *Let δ be a positive parameter. The notation of the preceding theorem is used. Then the congruent sets*

$$B_i(\delta) := \{q_i : q_i = u_i + K_i, \text{ any } K_i \text{ with } \|K_i\| < \delta\}, \quad (3.3)$$

$i = 1, \dots, k$, are an independent family if and only if

$$\delta < 1/\sqrt{k}. \quad (3.4)$$

Proof. Let $Q = (q_1, \dots, q_k)$ be the matrix of a collection of any vectors $q_i = u_i + K_i$ subject to (3.3). We have $Q = U + K$ where U is the unit-matrix and $K := (K_1, \dots, K_k)$. The q_i are linearly independent if and only if the determinant $|Q| \neq 0$, equivalently if and only if there does not exist a hyperplane $x'm = 0$ with some vector m satisfying $\|m\| = 1$, say, which contains all q_i . Now

$$\|Q'm\|^2 = 1 + 2m'K'm + \|K'm\|^2 \geq (1 - \|K'm\|)^2 \geq (1 - \delta\sqrt{k})^2 > 0,$$

the last inequality being true if $0 \leq \delta < 1/\sqrt{k}$; we have used the inequalities $|K'_i m| < \delta$ which imply $\|K'm\|^2 < k\delta^2$.

Necessity: Assume now the family $\{B_i(\delta), i = 1, \dots, k\}$ consists of independent sets for some value of $\delta \geq 0$. Then by definition any selection of k representative vectors q_i from them is linearly independent. This implies $Q'm \neq 0$ for all vectors m with $\|m\| = 1$, say. In order to show that necessarily $\delta < 1/\sqrt{k}$ choose the special vectors $K_i = \delta m$; this implies $\|K_i\| = \delta$ and

$$Q'm = (I + \delta um')m = m + \delta u, \quad u := (1, \dots, 1)'. \quad (3.5)$$

It follows

$$\|Q'm\|^2 = 1 + 2\delta m'u + \delta^2 k \quad \text{for all } m, \|m\| = 1. \quad (3.6)$$

The further specialization $m = -u/\sqrt{k}$ implies

$$\|Q'u\|^2/k = 1 - 2\delta\sqrt{k} + \delta^2 k = (1 - \delta\sqrt{k})^2 > 0, \quad (3.7)$$

where the latter inequality follows from the present assumption $Q'm \neq 0$ for all admissible δ -values; hence $\delta < 1/\sqrt{k}$ necessarily.

Statement (ii) of Theorem 3.1 may be proven by verifying that the hyperplane $x'u = 0$ contains the points $q_i = u_i - u/k$, $i = 1, \dots, k$ which in fact follows from $u'q_i = 1 - 1/k = 0$. (The proof could have been given also by verifying $\det Q = \det(I - uu'/k) = 0$.) \square

Remark 3.1

- (i) Any non-singular linear transformation B maps the congruent and open rotational cones D_i into independent elliptic and open cones with apices 0 and in general non-orthogonal axes; they are no longer congruent.
- (ii) Let $\tilde{K}_i \in R_k$, $\tilde{K}_i \perp u_i$, $\|\tilde{K}_i\| = 1$ and $\delta < 1/\sqrt{k-1}$. Then the vectors $\tilde{q}_i = u_i + \delta\tilde{K}_i$, $i = 1, \dots, k$ are linearly independent. Consequently D_i , $i = 1, \dots, k$ is an independent set family if the D_i are formed analogously to Theorem 3.1 but with the \tilde{K}_i instead of the K_i .

Theorem 3.3

(i) The spheres D'_i in (3.2) have the following $2^k - 1$ hyperplanes as tangent hyperplanes H_ϵ :

$$\{x \in R^k : x'\epsilon = 0\}, \quad \epsilon := (\epsilon_1, \dots, \epsilon_k)', \quad \epsilon_i = +1 \text{ or } -1, \quad (3.8)$$

$$\epsilon \neq (0, \dots, 0)'.$$

(ii) For a given separating tangent plane H_ϵ let $S := \{i : \epsilon_i > 0\}$. Then the open balls contained in D'_i with $i \in S$ lie on one side of the plane, those with $i \notin S$ lie on the other side.

The following corollary of this theorem concerns polyhedral cones.

Corollary 3.4. For $i = 1, \dots, k$ let Y_i be the polyhedral (pyramidal) cone spanned by the sphere D'_i and bordered by the $2^k - 1$ tangent and separating hyperplanes of the preceding Theorem 3.3. Then the open kernels of the Y_i form an independent family. Y_i and D'_i lie on the same side of any particular tangent hyperplane.

The proof is omitted because Theorem 2.3 is more general.

3.2. Rotational cones with unequal radii and non-orthogonal axes

Theorem 3.1 can be extended to rotational cones with unequal radii $\rho_i > 0$ respectively and non-orthogonal axes. First we discuss the case of unequal radii and orthogonal axes:

Theorem 3.5. Let $u_i = (0, \dots, 0, 1, 0, \dots, 0) \in R^k$ be the i th unit-vector and let K_i , $i = 1, \dots, k$, be any vectors. Then the open sets

$$D_i := \{q_i = u_i + K_i : \|K_i\| < \rho_i, \quad i = 1, \dots, k\}, \quad (3.9)$$

are an independent set family if

$$\sum_{i=1}^k \rho_i^2 < 1. \quad (3.10)$$

The same is true for the extended open rotational cones C_i in R^k generated by the sets D_i :

$$C_i = \{\lambda q_i : q_i \in D_i, \lambda \neq 0\}, \quad i = 1, \dots, k; \quad (3.11)$$

are an independent set family; the C_i have apices at the origin and central axes u_i .

Proof. Consider the hyperplane H defined by $n'x = 0$, $\|n\| = 1$. At contact points $u_i + K_i$ of H with the hyper-balls D_i , $i = 1, \dots, k$, the vectors K_i must be parallel to n . The vectors $u_i - \epsilon_i \rho_i n \in H$, $\epsilon_i = \pm 1$ represent contact points if the i th

component satisfies $n_i = \epsilon_i \rho_i$. This yields (3.10). (Warning: At this point perhaps a better notation for the unit-vectors u_i would be u^i .) \square

There are 2^{k-1} possibilities for the parameters ϵ_i corresponding to 2^{k-1} essential separating hyperplanes according to (2.3).

The case of rotational cones with unequal radii $\rho_i > 0$ respectively and non-orthogonal axes is handled by

Theorem 3.6. *Let $v_i \in R^k$ be any linearly independent vectors as axes of rotational cones with radii $\rho_i > 0$ respectively. Let $K_i, i = 1, \dots, k$, be any vectors. Instead of giving the matrix of the axes $V := (v_1, \dots, v_k)$ directly the equivalent parametric description $V = K\Lambda^{1/2}U$ will be used where the matrices U, K are orthonormal and $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_k)$ is the diagonal matrix of positive eigen-values in $UV'VU = \Lambda$; the diagonal matrix $\Lambda^{1/2}$ may have positiv or negativ elements. All these matrices are known. Then the open sets*

$$D_i := \{q_i = v_i + K_i : \|K_i\| < \rho_i, \quad i = 1, \dots, k\}, \quad (3.12)$$

are an independent set family if the numbers ρ_i satisfy

$$\sum_{i=1}^k \tilde{\rho}_i^2 / \lambda_i \leq 1; \quad (3.13)$$

here $U\rho = \tilde{\rho}$.

The straight-forward proof is omitted.

3.3. Families of independent polyhedral cones

Due to Theorem 2.3 there necessarily exists a half-space of R^k generated by $H_{\{1, \dots, k\}}$ that contains all open kernels of polyhedral cones of an independent set family. This fact implies the existence of an affine hyperplane parallel to the hyperplane $H_{\{1, \dots, k\}}$ which intersects all cones, and this fact can be used in constructions such as the following:

Denote by E the affine hyperplane $x_1 + \dots + x_k = 1$ spanned by the unit-vectors u_1, \dots, u_k of the R^k . Its coefficient vector is $u = (1, \dots, 1)'$, the diagonal vector of the hyper-octand $\cup_{i=1}^k \{x_i \geq 0\}$. The following construction can be carried over to any other diagonal vector $\epsilon := (\epsilon_1, \dots, \epsilon_k)'$, $\epsilon_j = \pm 1$ by a regular linear transformation. For $k \geq 1$ and any non-empty set $M \subset R^k$ we use the definitions of Section 2

$$\begin{aligned} \text{conv } M &:= \left\{ \sum_{i=1}^k \alpha_i a_i : a_i \in M, 0 \leq \alpha_i \text{ for all } i, \sum_{i=1}^k \alpha_i = 1 \right\}, \\ \text{cl } M &:= M \cup \{\text{cluster points of } M\}, \\ A &:= \text{conv } \{u_1, \dots, u_k\}, \end{aligned} \quad (3.14)$$

A being a $(k - 1)$ -dimensional simplex in E . Now at each vertex u_i of A we consider the ‘outer cone’ in E relative to A :

$$\begin{aligned}\kappa_i &:= \left\{ u_i + \sum_{r=1}^k t_r(u_i - u_r) : t_r \geq 0 \right\} \\ &= u_i + \text{pos} \{u_i - u_r : r = 1, \dots, k\},\end{aligned}\quad (3.15)$$

$i = 1, \dots, k$ (the term $t_i(u_i - u_i) = 0$ could have been left out). The sets κ_i are enlarged by the sets $S_i \subset E$ as follows:

$$\begin{aligned}S_i &:= \text{conv} \left\{ \kappa_i \cup \left\{ \frac{1}{2}(u_i + u_1), \dots, \frac{1}{2}(u_i + u_k) \right\} \right\} \\ &= \text{conv} \left\{ \kappa_i, \frac{1}{2}(u_i + u_1), \dots, \frac{1}{2}(u_i + u_k) \right\}\end{aligned}\quad (3.16)$$

$i = 1, \dots, k$ (the term $(u_i - u_i)/2 = u_i$ could again have been left out).

So S_i is a polyhedral set; it contains u_i in its relative interior. We construct next k -dimensional cones C_i in R^k (apex 0) using S_i as a $((k - 1)$ -dimensional) generating set:

$$C_i := \text{cl pos } S_i, \quad i = 1, \dots, k \quad (3.17)$$

(equivalently S_i can be considered as a ‘cross-cut’ and a ‘generating’ set of the cone C_i).

Remark 3.2

- (i) $\text{pos } S_i$ covers the interior of C_i but is not identical with C_i , therefore we take the (topological) closure of it.
- (ii) Using the operator

$$\text{pos}^* M := \{\alpha x : x \in M, 0 < \alpha\} \quad (3.18)$$

(note that this pinched cone does not contain the origin), we have

$$\text{pos}^* \kappa_i = \{2x_i u_i - x : x_r \geq 0 \text{ any } r, 2x_i > u'_i x\}, \quad x = (x_1, \dots, x_k)'. \quad (3.19)$$

The advantage of this construct is that it is very easy computationally to check whether or not some vector $y \in R^k$ belongs to the cone.

Lemma 3.7

- (i) The open kernels of the cones $C_i \subset R^k, i = 1, \dots, k$, defined by (3.17) form an independent set family. The C_i satisfy

$$C_i = \bigcap_{m=i}^k H_m^+ \quad \bigcap_{m=1}^{i-1} H_m^- \quad (3.20)$$

(up to linear transformations) with separating hyperplanes H_m defined by

$$x_1 + \cdots + x_m - x_{m+1} - \cdots - x_k = h'_S x = 0, \quad (3.21)$$

where $h_m := (u_1 + \cdots + u_m - u_{m+1} - \cdots - u_k)'$.

Hence the C_i are maximal with respect to this system of hyperplanes in the sense of Theorem 2.3(iv).

(ii) A fortiori the cones

$$\text{cl}(\text{pos } \kappa_i) \quad (3.22)$$

as well as the extended cones $\{C_i \cup (-C_i) : i = 1, \dots, k\}$ are independent set families.

Proof. We wish to apply Theorem 2.3. If we choose, for some m between 0 and k , any m of the cones C_i we may, up to interchanging the indices, assume them to be C_1, \dots, C_m . We claim: The hyperplane H_m according to (3.21) separates C_1, \dots, C_m from C_{m+1}, \dots, C_k in such a way that (without loss of generality) $C_i \subset H_m^+$ for $i \leq m$ and $C_i \subset H_m^-$ for $i > m$.

Since for any given m the operations cl , conv , pos do not lead out of H_m^+ or H_m^- respectively it suffices to show that κ_i and $(u_i + u_1)/2, \dots, (u_i + u_k)/2$ lie in H_m^+ for $i = 1, \dots, m$: We show first $\kappa_i \subset H_m^+$ (recalling $t_r \geq 0$, $r = 1, \dots, k$):

$$\begin{aligned} h'_m \left(u_i + \sum_{r=1}^k t_r (u_i - u_r) \right) \\ = h'_m u_i + t_1 [h'_m u_i - h'_m u_1] + \cdots + t_k [h'_m u_i - h'_m u_k] \\ = 1 + 2(t_{m+1} + \cdots + t_k) > 0. \end{aligned} \quad (3.23)$$

Analogously, we obtain for $i > m$

$$h'_m \left(u_i + \sum_{r=1}^k t_r (u_i - u_r) \right) = -1 - 2(t_1 + \cdots + t_m) < 0, \quad (3.24)$$

so that $\kappa_i \subset H_m^-$ for $i = m+1, \dots, k$.

Furthermore, for $i \leq m$ we have

$$h'_m((u_i - u_r)/2) = [h'_m u_i + h'_m u_r]/2 = \begin{cases} 1 & \text{for } r = 1, \dots, m \\ 0 & \text{for } r = m+1, \dots, k \end{cases}$$

and for $i > m$

$$h'_m((u_i - u_r)/2) = \begin{cases} 0 & \text{for } r = 1, \dots, m \\ -1 & \text{for } r = m+1, \dots, k. \end{cases} \quad (3.25)$$

We conclude $C_i \subset H_m^+$ for $i \leq m$ and $C_i \subset H_m^-$ for $i > m$.

The fact that equality holds in (3.20) is proved in the following lemma. This proves the first part of the lemma. The second part follows by applying central symmetry. \square

Without proof we cite

Lemma 3.8. *The cones C_i of the preceding lemma are maximal with respect to the set of separating hyperplanes H_m defined in (3.21) and their transforms; hence they satisfy (3.20).*

4. Results needed in the proof of Theorem 2.3

This section collects some lemmas and theorems are needed to prove Theorem 2.3, one of the main results of this paper.

Lemma 4.1. *Let \mathcal{L} be a linear space and $M_i, i = 1, \dots, m$ disjoint convex subsets of \mathcal{L} . Let $\text{conv}(\cup_i M_i)$ be the convex hull of $\cup_i M_i$. Then $x \in \text{conv}(\cup_i M_i)$ if and only if x is contained in an s -dimensional simplex Δ ($1 \leq s \leq m$) all whose vertexes lie in $\cup_i M_i$ but each M_i contains at most one vertex. Equivalently, x can be written as a convex combination of elements of M_i with at most one element from each M_i .*

The proof is based on the following version of Caratheodory's theorem (part of Theorem 1.23 [5] or Theorem 2.3 [3]):

Theorem 4.2. *Let \mathcal{L} be a linear space and $M_i, i = 1, \dots, m$ convex subsets of \mathcal{L} . Let $\text{conv}(\cup_i M_i)$ be the convex hull of $\cup_i M_i$. Then $x \in \text{conv}(\cup_i M_i)$ if and only if x is contained in an s -dimensional simplex Δ ($1 \leq s \leq m$) whose vertexes lie in $\cup_i M_i$.*

Proof (Proof of the preceding Lemma 4.1). The 'if'-direction of the assertion is obvious. For the other direction note that by the theorem cited above x is contained in a possibly degenerate simplex with m vertices $X_r \in \cup_i M_i, r = 1, \dots, m$:

$x = \sum_{r=1}^m \lambda_r X_r$ with $\lambda_r \geq 0, \sum_{r=1}^m \lambda_r = 1$. Define the index-sets $J_i = \{r : X_r \in M_i, 1 \leq r \leq m\}, i = 1, \dots, m$ (some J_i are possibly empty). Then

$$x = \sum_{i=1}^m \sum_{r \in J_i} \lambda_r X_r = \sum_{i=1}^m \left(\sum_{r \in J_i} \lambda_r \right) \sum_{r \in J_i} \lambda'_r X_r$$

where $\lambda'_r = 0$ if

$$\sum_{r \in J_i} \lambda_r = 0,$$

otherwise $\lambda'_r = \lambda_r (\sum_{r \in J_i} \lambda_r)^{-1} > 0$. Define the index-set

$$I = \left\{ i : \sum_{r \in J_i} \lambda_r > 0, \quad 1 \leq i \leq m \right\}. \quad (4.1)$$

Then for $i \in I$ we have $\sum_{r \in J_i} \lambda'_r = 1$. We put $Y_i = \sum_{r \in J_i} \lambda'_r X_r \in M_i$, the latter relation being true due to convexity of M_i . Moreover the Y_i are all different since the M_i are disjoint. It follows $x = \sum_{i \in I} \left(\sum_{r \in J_i} \lambda_r \right) Y_i$ which proves the lemma. Note that the dimension of the latter simplex whose interior contains x equals $|I|$. \square

The next result requires the Definition 4 of Section 2.2. The following separation theorem, a version of the Hahn–Banach theorem (e.g. Theorem 2.9 in [5]) is needed in the proof of Theorem 2.3:

Theorem 4.3 [5]. *If A and B are non-empty disjoint convex subsets in k -dimensional vector-space R^k then they can be separated by a hyperplane.*

References

- [1] F. Eicker, Distribution-free and robust linear regression with independent errors under most general assumptions, Contributed paper in Section 21, Conference Internat. Statistical Institute, Beijing, China, August 1995.
- [2] F. Eicker, Median estimates for robust and distribution-free estimation in multiple linear regression. An outline, in: J. Kunert, G. Trenkler (Eds.), Festschrift in honour of Siegfried Schach, entitled Mathematical statistics with applications in biometry, Eul, Lohmar; Köln, 2001.
- [3] G. Ewald, Combinatorial Convexity and Algebraic Geometry, Springer, New York, 1996.
- [4] J.G. Oxley, Matroid Theory, Oxford University Press, New York, 1992.
- [5] F.A. Valentine, Convex Sets, McGraw-Hill, New York, 1964.
- [6] R. Walter, Lineare Algebra und analytische Geometrie, Vieweg, Braunschweig, 1985.
- [7] G.M. Ziegler, oral communication.
- [8] G.M. Ziegler et al., Oriented Matroids, Encyclopedia of Mathematics 46, second ed., Cambridge University Press, Cambridge, 1999.